Gaussian target tracking with direction-of-arrival von Mises-Fisher measurements

Ángel F. García-Fernández, Filip Tronarp, Simo Särkkä, *Senior Member, IEEE*

***Abstract*—This paper proposes a novel algorithm for target tracking with direction-of-arrival measurements, modelled by von Mises-Fisher distributions. The algorithm makes uses of the assumed density framework with Gaussian distributions, in which the posterior probability density of the target state is approximated by a Gaussian density. A key component of this algorithm is that the proposed Bayesian model of the measurements takes into account the specific characteristics of angular measurements by using a von Mises-Fisher distribution. We propose two implementations of the algorithm, one based on first-order Taylor series expansion and another one based on sigma points. Simulation results show the benefits of the proposed algorithms in relation to other Gaussian filters in the literature.**

***Index Terms*—Bearings, direction-of-arrival, Kalman filtering, posterior linearisation, von Mises-Fisher distribution.**

I. Introduction

Target tracking and localisation using direction-of-arrival measurements is a central problem in many applications such as sonar, radar, and mobile positioning in long-term evolution (LTE) wireless communication systems [1]–[5]. In the Bayesian framework, this problem is usually posed by considering a target that moves according to a Markov process and is observed through noisy measurements. In this set-up, all information of interest about the target state at the current time step is contained in its probability density function (PDF) given all past and current measurements, which is referred to as posterior PDF. The posterior PDF can be expressed in a recursive manner by the Bayesian filtering recursion, which consists of the prediction and update steps [6].

The posterior is in general intractable, and therefore approximations are required. Particle filters can provide very accurate approximations to the posterior [7], though usually at a relatively high computational burden. Therefore, it is also of interest to develop computationally light algorithms, such as Gaussian filters, which propagate a Gaussian distribution through the filtering recursion. The most popular Gaussian filters for nonlinear systems are the extended Kalman filter (EKF) [1] and the sigma-point Gaussian filters such as the unscented Kalman filter (UKF) [8] or cubature Kalman filter (CKF) [9].

These Gaussian filters have been directly applied to directional measurements in a 2-D space, which can be represented by bearing angles [10]–[12], and in a 3-D space, which can be represented by azimuth and elevation angles [13], [14]. However, these methods do not take into account that angles have special characteristics. In particular, the use of averages and subtractions in Euclidean spaces for angles, which are required in the update step of the nonlinear Kalman filters, can be problematic [15], [16]. For example, the direct average of 170º and -170º is 0º, but it should be 180º [15]. The same happens for the difference, which is 340º, but it should be 20º. This becomes more problematic in a 3-D set-up in which bearings and elevation are being measured. One solution to avoid this problem is to use circular sums and subtractions instead of the Euclidean versions in the Kalman filter update, as explained in [17] for 2-D and 3-D cases, and also used in the simulations in [18] for a 2-D case. While this approach can work well, it is not a theoretically satisfying solution due to two reasons. First, a Gaussian density does not model the distribution of an angle well in general [15]. Second, the approach to compute the mean and variance of the angular measurement, which are required in the nonlinear Kalman filter update, is not consistent: the mean of the angular measurement is computed using the embedding approach, in which angles are represented as points on a unit sphere embedded in a Euclidean space [15], but the variance is computed using the wrapping approach [15].

Another way to process Gaussian bearing measurements that avoids averaging or subtracting angles is to perform a "pseudo-linearisation" of the measurement model [19], [20]. The pseudo-linearisation approach was extended to bearing/elevation measurements in [21], but with the additional assumption that the posterior PDF is independent in the z axis. In parallel to our work, a pseudo-linearisation method for bearing/elevation measurements without the previous assumption was proposed in [22].

In this paper, we propose the modelling of the direction-of-arrival measurements based on directional statistics [15], rather than Gaussian densities. To this end, we model the directional measurements using a von Mises-Fisher (VMF) distribution. The VMF distribution is a widely used distribution on the n-sphere that can be used to model bearing and azimuth/elevation measurements [15]. This approach has the benefit that the intrinsic characteristics of direction-of-arrival measurements are directly captured using its distribution. In addition, this modelling motivates the development of novel Gaussian filters that are tailored to this type of measurement model. In order to do so, this paper first explains how to compute the conditional mean and variance of the VMF measurements given the current target state and, based on them, how to perform statistical

A. F. García-Fernández is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, United Kingdom (email: angel.garcia-fernandez@liverpool.ac.uk). F. Tronarp and S. Särkkä are with the Department of Electrical Engineering and Automation, Aalto University, 02150 Espoo, Finland (emails: {filip.tronarp, simo.sarkka}@aalto.fi).
The authors would like to thank support from Academy of Finland projects 313708 and 314474.

linear regression (SLR) [23] for VMF measurements. With the SLR of VMF measurements, we can apply the iterated Gaussian filters in [24], which do not require additive Gaussian measurement noise. We propose two implementations of the algorithm. The first requires the calculation of the Jacobian of the measurement function, as in the EKF. The second one uses of sigma-points and transforms them through the measurement function to approximate the moments required for SLR, as in the UKF. However, there are important differences with the conventional EKF and UKF applied to angular data: the nonlinear measurement model is different, as we now consider a random VMF measurement lying in the n-sphere, which models directions. Moreover, there are additional correction factors to the resulting linearisations due to the VMF distribution of the measurements. An additional benefit of the proposed filters is that they deal in a unified way with 2-D and 3-D settings that involve directional measurements and can also be applied when there are additional non-directional measurements, for example, range.

The proposed algorithms are compared against other Gaussian filters in three numerical examples involving 2-D target tracking using bearings only measurements, 3-D target tracking using azimuth and elevation measurements, and 2-D target tracking using range/bearings measurements. On the whole, the proposed algorithms are the best performing ones.

The rest of the paper is organised as follows. Section II provides a background on the VMF distribution. The problem is formulated in Section III. The proposed Gaussian filters are developed in Section IV. A comparison between VMF and Gaussian models for angular data is provided in Section V. In Section VI, numerical simulations are presented. Finally, conclusions are drawn in Section VII.

II. Background on the von Mises-Fisher distribution

This paper deals with Bayesian filtering using directional measurements, which can correspond to bearing angles in 2-D or, azimuth and elevation angles in 3-D. Before dealing with Bayesian filtering, in this section, we first explain some concepts of directional statistics needed for understanding the rest of the paper.

In order to have an accurate mathematical model for a directional quantity, we must take into account its specific nature. A directional quantity z can be represented as a unit vector, so the underlying manifold structure should be accounted for in the model. We can express this property mathematically by writing $z \in S^{n-1}$ where $S^{n-1} = \left\{y:\ y^T y = 1,\, y \in \mathbb{R}^n\right\}$ is referred to as the $n-1$ sphere, which is a $n-1$-dimensional manifold that can be embedded in $\mathbb{R}^n$, and superscript T denotes transpose [15]. In practice, we usually consider $n = 2$, which represents the bearing of an object in a 2-D space, or $n = 3$, which represents the azimuth and elevation of an object in a 3-D space.

In order to perform Bayesian inference using angular variables, we need to consider distributions that characterise the knowledge we have about this type of random variable. In this paper, we will use von Mises-Fisher (VMF) distributions, which are parameterised by a mean direction μ and a concentration parameter κ, and are common unimodal distributions in S^{n-1}. The VMF density with respect to the uniform distribution is given by [15, Eq. (9.3.4)]

$$\mathcal{V}(z;\mu,\kappa) = \frac{\left(\frac{\kappa}{2}\right)^{n/2-1}}{\Gamma(n/2)\,\mathrm{I}_{n/2-1}(\kappa)} \exp\left(\kappa\mu^T z\right) \chi_{\|z\|=1}(z), \tag{1}$$

where $\kappa \geq 0$, $\|\mu\| = 1$, $\mathrm{I}_a(\cdot)$ represents the modified Bessel function of the first kind and order a, $\Gamma(\cdot)$ represents the gamma function and $\chi_A(\cdot)$ is the indicator function on set A. The higher κ is, the VMF distribution is more concentrated around μ. In addition, the VMF distribution is unimodal for $\kappa > 0$ and uniform for $\kappa = 0$. It should also be noted that for $n = 2$, the distribution reduces to the von Mises distribution on the circle [15], [25].

In this paper, we will make use of the mean and covariance matrix a VMF distribution. The expression for the mean can be found in [15] and for the second order moment in terms of the derivative of $A_n(\kappa)$ in [26, Eq. (5.6.2)]. We provide a proof with the specific equation of the covariance matrix we use in this paper in Appendix A. The mean and covariance matrix are then

$$\begin{aligned}\mathrm{E}[z] &= \int z\mathcal{V}(z;\mu,\kappa)\,dz \\ &= A_n(\kappa)\,\mu, \end{aligned} \tag{2}$$

$$\begin{aligned}\mathrm{C}[z] &= \int (z-\mathrm{E}[z])(z-\mathrm{E}[z])^T \mathcal{V}(z;\mu,\kappa)\,dz \\ &= \frac{A_n(\kappa)}{\kappa} I_n + \left[1 - A_n^2(\kappa) - n\frac{A_n(\kappa)}{\kappa}\right]\mu\mu^T, \end{aligned} \tag{3}$$

where I_n is the identity matrix of dimension n and

$$A_n(\kappa) = \frac{\mathrm{I}_{n/2}(\kappa)}{\mathrm{I}_{n/2-1}(\kappa)}, \tag{4}$$

where we remind that $\mathrm{I}_a(\cdot)$ is the modified Bessel function of the first kind and order a.

We would also like to note that the VMF distribution in 2-D can be parameterised by an angle by using polar coordinates to represent a point in the circle. In the same way, a VMF distribution in 3-D can be parameterised by two angles, azimuth and elevation, by using spherical coordinates. However, the computation of moments, such as mean and covariance, for angular/directional variables requires the parameterisation used in (1) [15]. Another advantage of this parameterisation is that we can deal with the 2-D and 3-D cases in a unified manner.

III. Problem formulation

In this paper, we are interested in performing Gaussian filtering with directional measurements using the assumed density framework. As the measurements only affect the update step, the prediction step is done as in the usual Gaussian filters, for instance, using closed-form formulas for linear/Gaussian dynamics, sigma-points as in the UKF, or using first-order Taylor series, as in the EKF, see [6] for more details. Therefore, we focus on the update as it is here

where novelty lies. We drop the time indices in the variables in Sections III to V for the sake of clarity, as all the variables refer to the time instant in which the update is performed.

The state of a target is represented as $x \in \mathbb{R}^{n_x}$ and has a prior density

$$p(x) = \mathcal{N}(x; \overline{x}, P),$$

where $\mathcal{N}(x; \overline{x}, P)$ denotes a Gaussian PDF with mean $\overline{x}$ and covariance matrix P evaluated at x.

We observe the target through a measurement vector $z = \left[(z_1)^T, ..., (z_m)^T\right]^T$ that consists of m angular measurements so $z_j \in S^{n-1}$ $j = 1, 2, ..., m$. We assume that given the target state, the angular measurements are independent so

$$p(z|x) = \prod_{j=1}^{m} p(z_j|x).$$

Then, we consider that each of the angular measurements follows a VMF distribution

$$p(z_j|x) = \mathcal{V}(z_j; h_j(x), \kappa_j) \tag{5}$$

where $h_j(\cdot)$ is a mapping from $\mathbb{R}^{n_x}$ to S^{n-1} that determines the mode and $\kappa_j \geq 0$ is the concentration parameter. Before continuing, we illustrate what the mapping $h_j(\cdot)$ looks like in a simple example.

Example 1. Let us consider the state represents a target moving in a 2D space. In this case, we often have $x = [p_x, v_x, p_y, v_y]^T$ where $[p_x, p_y]^T$ is the position vector and $[v_x, v_y]^T$ is the velocity vector. Then, if sensor j is located at $[s_{x,j}, s_{y,j}]^T$ and measures direction of arrival, then

$$h_j(x) = \frac{[p_x - s_{x,j}, p_y - s_{y,j}]^T}{\|[p_x - s_{x,j}, p_y - s_{y,j}]\|}. \square \tag{6}$$

In a Bayesian update, the objective is to compute the posterior PDF, which is the PDF of the state given the measurement, and is given by Bayes' rule

$$p(x|z) \propto p(z|x)\, p(x), \tag{7}$$

where $\propto$ denotes proportionality.

The posterior (7) is intractable to calculate for the considered measurement model and, in this paper, the objective is to obtain a Gaussian approximation to this posterior PDF.

IV. Gaussian filtering with direction-of-arrival VMF measurements

In this section, we explain how we perform Gaussian filtering with directional measurements modelled by VMF distributions.

A. Gaussian filtering based on linearisation

In this section, we explain the proposed algorithm to obtain a Gaussian approximation to the posterior (7). For a given x, we can write that the relation between z_j and x, which is given by (5), as a measurement equation of the form

$$z_j = g_j(x) + \eta_j(x), \tag{8}$$

where $g_j(x) = \mathrm{E}[z_j|x]$, which is a nonlinear transformation of x, and $\eta_j(x)$ is a zero-mean noise with covariance matrix $R_j(x) = \mathrm{C}[z_j|x]$ conditioned on x. It should be noted that $\eta_j(x) = z_j - g_j(x)$ is uncorrelated with x and $g_j(x)$, which means that $\mathrm{C}[x, \eta_j(x)] = 0$ and $\mathrm{C}[g_j(x), \eta_j(x)] = 0$ [27]. Also, $g_j(x)$ is a vector of dimension n and $R_j(x)$ a matrix of dimensions $n \times n$.

We have highlighted with the notation that z_j can be seen as a nonlinear transformation $g_j(\cdot)$ of x plus an uncorrelated zero-mean noise with a covariance matrix $R_j(x)$ that depends on x. The conditional moments $g_j(x) = \mathrm{E}[z_j|x]$ and $R_j(x) = \mathrm{C}[z_j|x]$ can be directly obtained from (2), (3) and (5) as

$$g_j(x) = A_n(\kappa_j)\, h_j(x), \tag{9}$$

$$R_j(x) = \frac{A_n(\kappa_j)}{\kappa_j} I_n + \left[1 - A_n^2(\kappa_j) - n\frac{A_n(\kappa_j)}{\kappa_j}\right] h_j(x)\, h_j^T(x). \tag{10}$$

A widely-used approach to provide Gaussian approximations to the posterior when the measurement equation is non/linear-non/Gaussian, such as (8) is to linearise it as [24], [28]

$$z_j \approx A_j x + b_j + r_j,$$

where $A_j \in \mathbb{R}^{n \times n_x}$, $b_j \in \mathbb{R}^n$ and r_j is a zero-mean noise with covariance matrix $\Omega_j \in \mathbb{R}^{n \times n}$, uncorrelated with x.

For this kind of system, the linear mean square error (LMMSE) estimator and its mean square error matrix are available in closed-form and can be used to approximate the first two moments of the posterior [27]. Under the additional assumption that r_j is Gaussian, the posterior becomes Gaussian with mean and covariance matrix

$$\overline{u} = \overline{x} + PA^T\left(APA^T + \Omega\right)^{-1}(z - A\overline{x} - b), \tag{11}$$

$$W = P - PA^T\left(APA^T + \Omega\right)^{-1}AP, \tag{12}$$

where $A = \left[A_1^T, ..., A_m^T\right]^T$, $b = \left[b_1^T, ..., b_m^T\right]^T$ and $\Omega = \mathrm{diag}(\Omega_1, ..., \Omega_m)$. Notably, the accuracy of the posterior approximation only depends on how we choose A, b and Ω. We proceed to explain how to select these parameters.

B. Statistical linear regression of VMF measurements

In this section, we review the concept of statistical linear regression, which is important to select the linearisation parameters. In its general setting, statistical linear regression [24] requires a density $p(\cdot)$ on variable x, whose first two moments are $\overline{x}$ and P, and another random variable z_j that has a conditional density $p(z_j|x)$. We can always write z_j as a nonlinear transformation $g_j(x)$ of x plus uncorrelated zero-mean noise $\eta_j(x)$, see (8). In SLR, we select A_j^+ and b_j^+ that minimise the mean square error of $g_j(x) + \eta_j(x)$ and its linear approximation so that

$$\left(A_j^+, b_j^+\right) = \underset{(A_j, b_j)}{\arg\min}\, \mathrm{E}\left[\|g_j(x) + \eta_j(x) - A_j x - b_j\|^2\right] \tag{13}$$

$$= \underset{(A_j, b_j)}{\arg\min}\, \mathrm{E}\left[\|g_j(x) - A_j x - b_j\|^2\right], \tag{14}$$

where the expectation is taken with respect to $p(\cdot)$ and $\|a\|^2 = a^T a$. Also, we would like to remark that the step from (13) to (14) holds because $g_j(x)$ and $\eta_j(x)$ are uncorrelated [27]. The solution to (14) is given by

$$A_j^+ = \mathrm{C}\left[x, g_j(x)\right]^T P^{-1}, \tag{15}$$

$$b_j^+ = \mathrm{E}\left[g_j(x)\right] - A_j^+ \overline{x}. \tag{16}$$

The resulting mean square error matrix is

$$\begin{aligned}\Omega_j^+ &= \mathrm{E}\left[\left(g_j(x) - A_j^+ x - b_j^+\right)\left(g_j(x) - A_j^+ x - b_j^+\right)^T\right] \\ &\quad + \mathrm{E}\left[\eta_j(x)\,\eta_j^T(x)\right] \\ &= \mathrm{C}\left[g_j(x)\right] + \mathrm{E}\left[R_j(x)\right] - A_j^+ P \left(A_j^+\right)^T. \end{aligned} \tag{17}$$

The moments found in (15)-(17) can be written more explicitly in terms of moments $h_j(\cdot)$ by using (9) and (10), which yields

$$\mathrm{E}\left[g_j(x)\right] = A_n(\kappa_j)\,\mathrm{E}\left[h_j(x)\right], \tag{18}$$

$$\mathrm{C}\left[x, g_j(x)\right] = A_n(\kappa_j)\,\mathrm{C}\left[x, h_j(x)\right], \tag{19}$$

$$\mathrm{C}\left[g_j(x)\right] = \left(A_n(\kappa_j)\right)^2 \mathrm{C}\left[h_j(x)\right], \tag{20}$$

$$\begin{aligned}\mathrm{E}\left[R_j(x)\right] &= \frac{A_n(\kappa_j)}{\kappa_j} I_n + \left[1 - A_n^2(\kappa_j) - n\frac{A_n(\kappa_j)}{\kappa_j}\right] \\ &\quad \times \mathrm{E}\left[h_j(x)\,h_j^T(x)\right]. \end{aligned} \tag{21}$$

The moments in (18)-(21) do not have closed-form expressions so they must be approximated. We proceed to discuss two approximations.

1) Sigma-point approximation: In the sigma-point approximation of the SLR, the moments (18)-(21) are approximated using a sigma-point method, such as the unscented transform [6]. Moments (18)-(21) can be written in terms of the moments of $h_j(\cdot)$, which is a mapping from the state-space to the unit sphere, representing the directional measurement. In a typical Gaussian filtering scenario with real measurements and additive measurement noise, we use the moments $\mathrm{E}\left[h_j(x)\right]$, $\mathrm{C}\left[x, h_j(x)\right]$ and $\mathrm{C}\left[h_j(x)\right]$ to perform the SLR and compute the Gaussian approximation to the posterior [23], [28]. The interesting aspect of the SLR applied to VMF angular measurements is that, rather than using the moments of $h_j(\cdot)$ directly to obtain the SLR, they should be modified according to the formulas in (18)-(21), to then compute (15)-(17). Then, the sigma-point approximation to the SLR for VMF measurements first select m_s sigma-points $\mathcal{X}_1, ..., \mathcal{X}_{m_s}$ and weights $\omega_1, ..., \omega_{m_s}$, which match the moments $\overline{x}$ and P [6], [8], [9]. Then, the moments of $h_j(\cdot)$ and the SLR can be calculated as indicated in Algorithm 1.

2) First order Taylor series approximation: Another approach to approximate SLR is to use a first order Taylor series expansion of the conditional moments $g_j(x) = \mathrm{E}\left[z_j|x\right]$ and $R_j(x) = \mathrm{C}\left[z_j|x\right]$ around the prior mean $\overline{x}$ [24]. In this case, the expectation of the first order term in the Taylor series expansion of $R_j(x)$ vanishes when computing $\mathrm{E}\left[R_j(x)\right]$. Therefore, we just need to consider the approximations

$$\begin{aligned} g_j(x) &\approx g_j(\overline{x}) + \nabla g_j(\overline{x})(x - \overline{x}), \\ R_j(x) &\approx R_j(\overline{x}), \end{aligned}$$

Algorithm 1 Statistical linear regression of a VMF measurement using sigma-points

Input: Function $h_j(\cdot)$, concentration parameter κ_j, see (5), and the first two moments $\overline{x}$, P of a PDF $p(\cdot)$.
Output: SLR parameters $\left(A_j^+, b_j^+, \Omega_j^+\right)$.
- Select m_s sigma-points $\mathcal{X}_1, \ldots, \mathcal{X}_{m_s}$ and weights $\omega_1, \ldots, \omega_{m_s}$ according to $\overline{x}$ and P [6].
- Transform the sigma-points $\mathcal{H}_i = h(\mathcal{X}_i) \quad i = 1, \ldots, m_s$.
- Approximate the moments

$$\mathrm{E}\left[h_j(x)\right] \approx \sum_{i=1}^{m_s} \omega_i \mathcal{H}_i,$$

$$\mathrm{C}\left[x, h_j(x)\right] \approx \sum_{i=1}^{m_s} \omega_i \left(\mathcal{X}_i - \overline{x}\right)\left(\mathcal{H}_i - \mathrm{E}\left[h_j(x)\right]\right)^T,$$

$$\mathrm{E}\left[h_j(x)\,h_j^T(x)\right] \approx \sum_{i=1}^{m_s} \omega_i \mathcal{H}_i \mathcal{H}_i^T,$$

$$\mathrm{C}\left[h_j(x)\right] = \mathrm{E}\left[h_j(x)\,h_j^T(x)\right] - \mathrm{E}\left[h_j(x)\right]\left(\mathrm{E}\left[h_j(x)\right]\right)^T.$$

- Obtain A_j^+, b_j^+, Ω_j^+ using (18)-(21) and (15)-(17).

where we recall that $g_j(\cdot)$ and $R_j(\cdot)$ are given by (9) and (10), and therefore $\nabla g_j(\overline{x}) = A_n(\kappa_j)\,\nabla h_j(\overline{x})$ with $\nabla h_j(\overline{x})$ denoting the Jacobian of $h_j(\cdot)$ evaluated at $\overline{x}$. The resulting steps for the SLR approximation are provided in Algorithm 2. It should be noted that, in this case, the SLR approximation of $\left(A_j^+, b_j^+, \Omega_j^+\right)$ does not take P into account and that $\Omega_j^+ = R_j(\overline{x})$.

Algorithm 2 Statistical linear regression of a VMF measurement using first-order Taylor series of $g_j(\cdot)$ and $R_j(\cdot)$

Input: Function $h_j(\cdot)$, Jacobian $\nabla h_j(\cdot)$, concentration parameter κ_j, see (5), and the first two moments $\overline{x}$, P of a PDF $p(\cdot)$.
Output: SLR parameters $\left(A_j^+, b_j^+, \Omega_j^+\right)$.
- $\mathrm{E}\left[g_j(x)\right] = A_n(\kappa_j)\,h_j(\overline{x})$.
- $\mathrm{C}\left[x, g_j(x)\right] = A_n(\kappa_j)\,P\left(\nabla h_j(\overline{x})\right)^T$.
- $\mathrm{C}\left[g_j(x)\right] = A_n^2(\kappa_j)\,\nabla h_j(\overline{x})\,P\left(\nabla h_j(\overline{x})\right)^T$.
- Obtain $\mathrm{E}\left[R_j(x)\right] \approx R_j(\overline{x})$ using (10).
- Obtain A_j^+, b_j^+, Ω_j^+ using (18)-(21) and (15)-(17).

3) Use of SLR in the update step: A common approach to perform the Gaussian filtering update is to use (11)-(12) with the parameters A, b, Ω chosen by SLR of the measurement function with respect to the prior, whose first two moments are $\overline{x}$ and P, approximated with a first-order Taylor series or sigma-points [23], [28]. The same approach can be used for angular measurements: we run Algorithms 1 or 2 with input $\overline{x}$, P and then we perform the update with the resulting values of A, b, Ω. We should note that, in order to perform this update, we only need the inverses of P and $APA^T + \Omega$, which are positive-definite matrices.

The main difference between the proposed approach and the usual EKF and UKF is that the EKF and UKF were not designed to deal with the specific properties of directional measurements, which lie in a manifold (see Section II). EKF and UKF usually model directions as angles with additive Gaussian noise and deal with angles as if they were numbers in the real line. This implies that the chosen values of A, b, Ω for the EKF and UKF are not necessarily suitable, especially, if

there are angles near the π radians boundary. On the contrary, by proper modelling of the directional measurements as a VMF distributions, which are related to Gaussian distributions (see Section V), we avoid these problems and are able to improve performance.

It is also worth noting that Gaussian filters that perform SLR with respect to the prior, applied to measurements with additive Gaussian noise, provide an accurate approximation to the posterior for high measurement noise in relation to the mean square error matrix, see (17). This was proved in [29] by calculating the Kullback-Leibler divergence (KLD) between the true density of the state and the measurement and its approximation [29]. The KLD analysis in [29] is not applicable to the case of VMF measurements, as it is restricted to measurements with additive Gaussian noise. Nevertheless, we show in Appendix B that, if $\kappa_j \to 0$ for $j \in \{1, ..., m\}$, the posterior and the approximated posterior by the proposed filters tend to the prior. In other cases, the posterior approximation is not guaranteed to be accurate, so one can benefit from iterated filters that look for an optimal linearisation, as we proceed to explain next.

C. Iterated statistical linear regression

As indicated in [24], [28], we can improve estimation by performing iterated SLR with respect to the current approximation of the posterior. We proceed to explain this idea.

Performing SLR with respect to the prior provides us with the optimal A_j^+ and b_j^+ in the sense of minimising the mean square error in (14). However, this minimisation does not take into account the observed value of the measurement z_j, which we know. Therefore, in order to take all available knowledge into account, the best affine approximation of the nonlinear function $g_j(\cdot)$ is given by

$$\left(A_j^+, b_j^+\right) = \underset{(A_j, b_j)}{\arg\min} \, \mathrm{E}\left[\left\|g_j(x) - A_j x - b_j\right\|^2 | z\right].$$

The solution is then given by SLR of $g_j(\cdot)$ with respect to the posterior, not the prior. We also set $\Omega_j^+ = \mathrm{E}\left[\left(g_j(x) - A_j^+ x - b_j^+\right)(\cdot)^T | z\right] + \mathrm{E}\left[\eta_j(x)(\cdot)^T | z\right]$, which is the covariance matrix of the residual $g_j(x) - A_j^+ x - b_j^+$ plus the covariance matrix of $\eta_j(x)$ in the area indicated by the posterior. We cannot compute the expectations with respect to the posterior, as the posterior is what we aim to calculate. Nevertheless, this idea leads to an iterated scheme, which is referred to as iterated posterior linearisation filter (IPLF).

We start with the prior moments $\overline{u}^1 = \overline{x}$, $W^1 = P$ and calculate A_j^1, b_j^1, Ω_j^1 : using the SLR of the j angular measurement with respect to $\overline{u}^1$, W^1. Based on A_j^1, b_j^1, Ω_j^1, we use (11) and (12) to obtain $\overline{u}^2$ and W^2, which characterise the posterior approximation at the second iteration. The iterations continue until a convergence criterion is met or for a fixed number n_{it} of steps. The resulting algorithm is given in Algorithm 3. It should be noted that, when we apply (11) and (12) for different iterations, only A, b and Ω change, the prior moments $\overline{x}$ and P remain unchanged.

We would also like to note that, to make our main points clearer, we have presented the update step of the algorithm when there are only angular measurements. Nevertheless, if there are additional non-angular measurements, such as range, they can be directly included by performing the corresponding SLR. That is, for measurements with additive Gaussian noise, we perform SLR of the measurement function, see [28]. We analyse a scenario of this type that includes bearings and also range measurements in Section VI-B.

Algorithm 3 The update step of the IPLF with direction-of-arrival VMF measurements

Input: Prior moments $\overline{u}^1 = \overline{x}$, $W^1 = P$, measurement z_j and model parameters $(h_j(\cdot), \kappa_j)$ for $j = 1, ..., m$, and number n_{it} of iterations.
Output: Posterior moments $\overline{u}^{n_{it}+1}$, $W^{n_{it}+1}$.

```
for i = 1 to n_it do
    for j = 1 to m do
        - Obtain A_j^i, b_j^i, Ω_j^i :
            - Run Algorithm 1, for sigma-point implementation, or Algorithm 2, for first-order Taylor series implementation, with input ū^i, W^i, h_j(·) and κ_j.
    end for
    - Compute ū^i, W^i using (11) and (12) with
        - A = [(A_1^i)^T, ...., (A_m^i)^T]^T, b = [(b_1^i)^T, ...., (b_m^i)^T]^T,
        - Ω = diag(Ω_1^i, ..., Ω_m^i).
end for
```

In addition, we would like to mention that for high values of κ, there can be numerical problems to calculate $A_n(\kappa)$ directly using (4). In this case, it is more suitable to use the asymptotic expansion of $A_n(\kappa)$ for large κ [15, Eq. (10.3.6)] and make the approximation

$$A_n(\kappa) \approx 1 - \frac{n-1}{2\kappa}. \tag{22}$$

This implies that, for large κ, the eigenvalue of $R_j(x)$ for the eigenvector in the direction of $h_j(x)$ has order $O\left(1/\kappa_j^2\right)$ and the eigenvalues of $R_j(x)$ for the eigenvectors in the orthogonal directions of $h_j(x)$ are all equal and have order $O(1/\kappa_j)$, which implies that $R_j(x)$ becomes close to singular. Nevertheless, the filter can run properly for large large κ_j just by using (22) in Algorithm 3.

V. Comparison of von Mises-Fisher and Gaussian measurement models

In this section, we discuss the differences between the VMF model for directional measurements used in this paper and the usual Gaussian additive models. We address the 2-D and 3-D cases separately as relevant differences arise.

A. 2-D case

Let us consider a target in a two dimensional space. A bearing measurement is usually modelled as

$$z_\varphi = \mathrm{atan2}\left(p_y - s_{y,j}, p_x - s_{x,j}\right) + \zeta_\varphi, \tag{23}$$

where $[p_x, p_y]^T$ denotes the target position, $[s_{x,j}, s_{y,j}]^T$ denotes the sensor position, ζ_φ is a zero-mean Gaussian noise with variance R_φ, and $\mathrm{atan2}(\cdot,\cdot)$ is the four-quadrant inverse tangent. We have already mentioned that (23) does not model angular measurements well in general, as it does not take into account that measurements z_φ and $z_\varphi + 2\pi k$, for any integer

k, are equivalent. Nevertheless, we can establish a connection with the von Mises measurement model given by (5) and (6), with $p = 2$.

We can represent a von Mises density in terms of the angle using polar coordinates $z_j = [\cos z_\varphi, \sin z_\varphi]^T$, $h_j(x) = [\cos \mu_\varphi, \sin \mu_\varphi]^T$, where $\mu_\varphi = \operatorname{atan2}(p_y - s_{y,j}, p_x - s_{x,j})$. The von Mises density and the wrapped Gaussian density are quite similar if they are matched to have the same first trigonometric moment [15], but model (23) considers standard Gaussian densities, not wrapped ones. In this case, if κ_j is high, the VMF distribution tends to a Gaussian distribution with mean μ_φ and variance $1/\kappa_j$ [15, Sec. 3.5.4]. However, even for large κ_j, nonlinear Kalman filters applied with the Gaussian model (23) require ad-hoc modifications [17] if the density of the predicted measurement is non-negligible around π radians. Nevertheless, we prove in Appendix C that if the density of the predicted angular measurement is narrowly concentrated around a point and is negligible around π radians boundary, the proposed non-iterated filter and a nonlinear Kalman filter applied to measurement (23) are approximately equal.

B. 3-D case

Let us consider a target in a three dimensional space. Let its position and the sensor position be $[p_x, p_y, p_z]^T$ and $[s_{x,j}, s_{y,j}, s_{z,j}]^T$, respectively. A typical model for azimuth-elevation measurements is to consider a bearing/azimuth measurement as in (23) and an elevation measurement [17]

$$z_\theta = \arcsin\left(\frac{p_z - s_{z,j}}{\|[p_x - s_{x,j},\, p_y - s_{y,j},\, p_z - s_{z,j}]\|}\right) + \zeta_\theta, \tag{24}$$

where ζ_θ is a zero-mean Gaussian noise with variance R_θ, which is independent of the bearing noise. Even though this model does not take into account the specific properties of aximuth/elevation measurements, one advantage of this model with respect to the VMF model is that it can consider different noise variances for azimuth/elevation, which can be useful in some cases. However, there are other drawbacks that we proceed to explain.

In contrast to the 2-D case, a VMF measurement on the sphere does not have an analogue with azimuth/elevation measurements with additive Gaussian noise, as in (23)-(24). In order to illustrate this, we consider two VMF distributions with the same concentration parameter 500 and modes located at azimuth 0º/elevation 0º and azimuth 0º/elevation 75º, respectively. We show 10^4 samples from these VMF distributions on the sphere and with azimuth/elevation representation, which is obtained by using spherical coordinates, in Figure 1. For elevation 0º, the corresponding azimuth/elevation samples can be approximated as Gaussian with covariance matrix $I_2/500$. However, for elevation 75º, the samples show more spread on the azimuth. This is due to the fact that the closer we are to the poles (elevation close to 90º or -90º), the meridians are also closer. Therefore, a azimuth/elevation sensor that provides the same uncertainty about the target in all directions cannot be accurately modelled using (23)-(24). Model (23)-(24) implies that the same azimuth uncertainty translates in

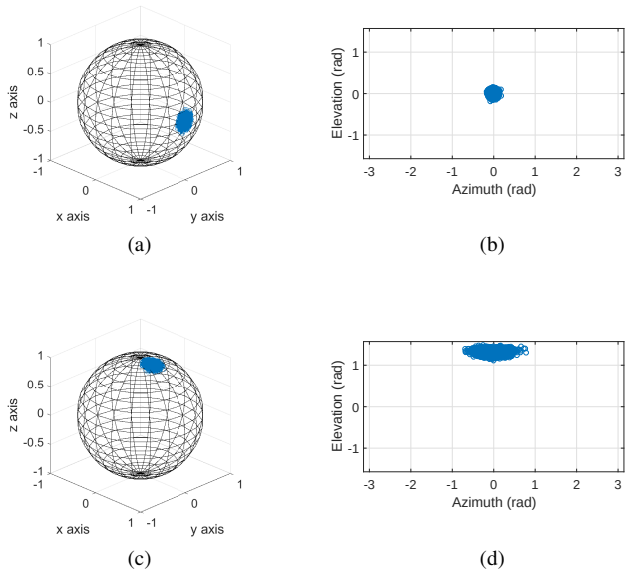

Figure 1: Plot of 10^4 samples of a VMF distribution with concentration parameter 500 on the unit sphere and with azimuth/elevation representation: (a) and (b) with a mode with azimuth 0º and elevation 0º and (c) and (d) with a mode with azimuth 0º and elevation 75º. The higher the elevation, the samples become more spread in azimuth as meridians are closer near the poles.

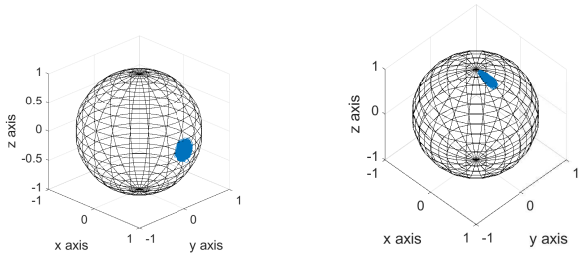

Figure 2: Plot of 10^4 samples on the unit sphere of a Gaussian distribution on the bearing/elevation with mean 0º/0º (left) and 0º/75º (right) and covariance matrix $1/500 I_2$. Samples are more concentrated in the bearing direction as elevation increases.

lower uncertainty regarding target location for elevations near the poles. For example, for $R_\varphi = R_\theta = 1/500$, and mean located at azimuth 0º/elevation 0º, 10^4 samples from (23)-(24) look as in Figure 1(a) and (b). However, if the mean is located at azimuth 0º/elevation 75º, the samples on the sphere have a different pattern with less uncertainty along the meridians, see Figure 2. This property does not model well an azimuth/elevation sensor based on an antenna whose beam direction is controlled mechanically or electronically [30]. A VMF distribution is more suitable.

In addition, as happens in the 2-D case, nonlinear Kalman filters applied with model (23)-(24) require ad-hoc modifications [17] if the density of the predicted measurement is non-negligible near the poles or the bearing boundary at π radians.

VI. Simulations

In this section, we compare the performance of the developed algorithms with relevant Gaussian filters for direction-

of-arrival measurements in three target tracking scenarios. The scenario in Section VI-A considers 2-D target tracking with bearings only measurements. The scenario in Section VI-B considers 2-D target tracking with range-bearings measurements. Finally, the scenario in Section VI-C deals with 3-D target tracking with azimuth and elevation measurements.

The simulations have been run in Matlab and we have used the codes in [31] and [32] to generate the VMF samples in 2-D and 3-D cases, respectively. In this section, all the units are given in the international system.

A. Bearings only tracking scenario

We consider a scenario in which a 2-D target is tracked based on angle-of-arrival measurements of three sensors. The state vector at time k is $x^k = \left[p_x^k, \dot{p}_x^k, p_y^k, \dot{p}_y^k\right]^T$, which includes position and velocity. The dynamic model of the target is the nearly-constant velocity model:

$$x^{k+1} = Fx^k + v^k, \tag{25}$$

$$F = I_2 \otimes \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix}, \tag{26}$$

where $\otimes$ is the Kronecker product and v^k is the process noise at time k. The process noise is zero-mean Gaussian distributed with covariance matrix

$$Q = qI_2 \otimes \begin{pmatrix} \tau^3/3 & \tau^2/2 \\ \tau^2/2 & \tau \end{pmatrix}, \tag{27}$$

where q is a parameter of the model. The prior at time 0 is

$$p(x^0) = \mathcal{N}\left(x^0; \bar{x}^0, \Sigma^0\right), \tag{28}$$

where $\bar{x}^0 = [-100, 7, 0, 5]^T$ and $\Sigma^0 = \text{diag}\left(\left[20^2, 1, 1, 1\right]\right)$. In the simulations, we consider $q = 0.25$ and $\tau = 0.5$.

The target is observed by three angle-of-arrival sensors located at $[s_{x,1}, s_{y,1}]^T = [100, 0]^T$, $[s_{x,2}, s_{y,2}]^T = [0, -100]^T$, and $[s_{x,3}, s_{y,3}]^T = [0, 150]^T$. Sensor j measures bearing angle every 3 time steps starting at time step j. Therefore, at each time step, only one of the sensors takes a measurement. We first consider measurements that are distributed according to a VMF distribution (5) with measurement function (6) and concentration parameter $\kappa_j = 200$ for $j = \{1, 2, 3\}$.

We have implemented the VMF-IPLF developed in this paper with 1 (VMF-IPLF1) and 5 (VMF-IPLF5) iterations, using the sigma-point and the first-order Taylor series implementations. The first-order Taylor series implementations are referred to as VMF-IPLF1(T) and VMF-IPLF5(T). In addition, we have implemented the conventional extended KF (EKF) and unscented KF (UKF) [6], the UKF with circular sums and subtractions for the measurement [17], which is referred to as angular UKF (AUKF), and the filters described in [20]: the pseudolinear KF (PLKF), the bias compensated pseudolinear KF (BC-PLKF), and instrumental variable-based KF (IVKF). The UT for the UKF, the AUKF and the VMF-IPLF has been implemented with $m_s = 2n_x + 1$ sigma-points and the weight of the sigma-point located on the mean is 1/3. The filters that require an additive Gaussian measurement model (EKF, UKF, AUKF, PLKF, BC-PLKF, and IVKF) consider a measurement model as in (23) with a zero-mean Gaussian noise with variance $1/\kappa_j$, which is an accurate approximation of (5) for large κ_j [15]. The IVKF has been implemented with $\kappa = 2$ (κ as defined in [20, Sec. V]). We have also modified the IVKF by substituting the direct subtraction of angles in [20, Eq. (34)] by circular subtraction [17], as it improves performance.

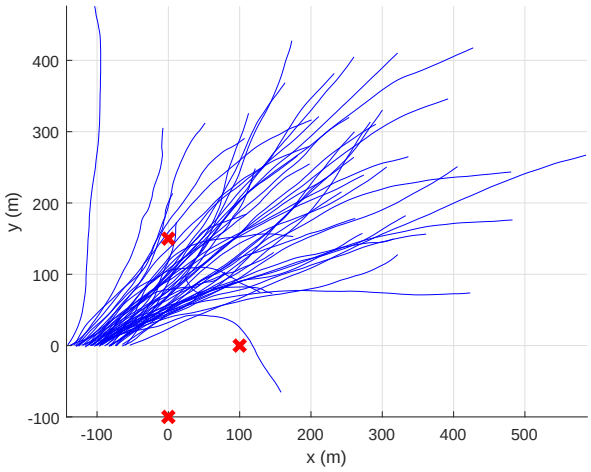

Figure 3: 2-D bearings only tracking scenario. The considered 50 target trajectories are shown in blue. The three sensor positions are marked as red crosses.

We evaluate the algorithm with $N_{mc} = 10^3$ Monte Carlo runs with 50 different trajectories of length $K = 100$ drawn from the dynamic model. Measurements are sampled from the VMF model. The considered trajectories and the scenario of the simulations are shown in Figure 3. This scenario is challenging for the filters as the target is observed at different angles and distances from the sensor that takes the measurement. In this scenario, the conventional EKF and UKF perform quite badly, as some sensors measure bearing near π radians at some time steps. The PLKF and BC-PLKF do not work well in this scenario either. Due to their bad performance, all these filters are not considered further.

The root mean square (RMS) error for the position elements provided by the rest of the filters is shown in Figure 4. We can see that up to time step 50 all the filters behave similarly well. From this time step, all the variants of the proposed algorithm VMF-IPLF outperform IVKF and AUKF. In terms of RMS error, the iterations barely affect estimation performance in this scenario. It should be noted that, at the end of the simulation, the target is moving away from the sensors, see Figure 3. Therefore, the RMS error increases towards the end of the simulation as the measurements provided by the sensors are less informative to locate targets at far distances.

We also show the normalised estimation error squared (NEES) [1] of the position to check the consistencies of the filters in Figure 5. We can see that the iterations of the VMF-IPLF improve consistency, especially between time steps 40 and 60. VMF-IPLF1(T) shows inconsistent behaviour at time step 47 but it soon recovers. Importantly, increasing the iterations to 5 markedly improves consistency. Also, AUKF has slightly worse consistency than VMF-IPLF5 in the time

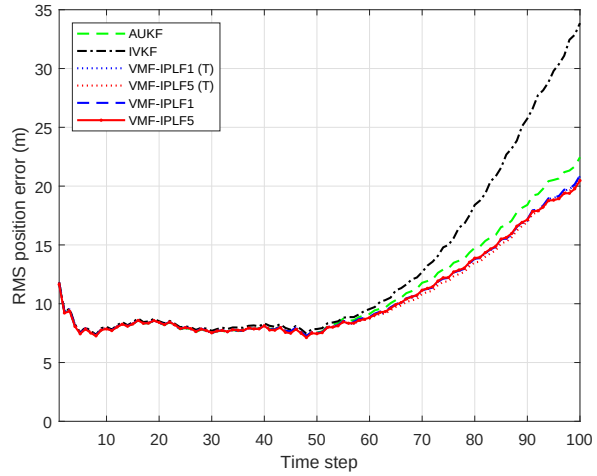

Figure 4: RMS position error against time for the 2-D bearings-only tracking scenario. The lines of all variants of the VMF_IPLF overlap.

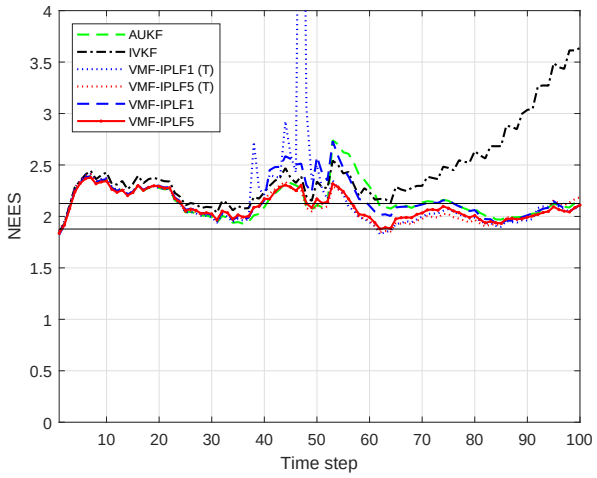

Figure 5: NEES of the position for the 2-D bearings-only tracking scenario. Two horizontal black lines indicate the 95% probability region.

window between 40 and 60. IVKF is the worst performing filter in RMS error and NEES. The execution times in milliseconds of our Matlab implementations of the algorithms on a computer with a 3.5 GHz Intel Xeon E5 processor are shown in Table I. On the whole, filters that perform better use more computational resources, but all of them are require low computational burden.

We proceed to analyse the performance of the filters with respect to different concentration parameters and the underlying distribution of the measurement. We consider $\kappa_j \in \{100, 200, 300\}$, which is equal for all sensors. We also consider experiments with measurement samples drawn from a VMF distribution and drawn from the Gaussian additive model that corresponds to the VMF distribution, which was explained above. We analyse the results by the RMS error for the position elements considering all time steps and normalised by the number of time steps. This error is defined as

$$E \triangleq \sqrt{\frac{1}{KN_{mc}} \sum_{l=1}^{N_{mc}} \sum_{k=1}^{K} \left(\hat{p}_k^l - p_k\right)^T \left(\hat{p}_k^l - p_k\right)}, \qquad (29)$$

where p_k denotes the true position of the target at time k and $\hat{p}_k^j$ its estimation at the lth Monte Carlo run.

Table II shows error (29) for the different algorithms, concentration parameters and true distribution of the measurement samples. In general, the VMF-IPLF performs slightly better than AUKF for VMF and Gaussian measurements. In this scenario, the Taylor series implementation of VMF-IPLF generally works better than the sigma-point implementation in terms of RMS. IVKF performs worse than AUKF and VMF-IPLF, especially, as the concentration parameter gets smaller.

Table I: Computational times in milliseconds of the algorithms in the bearings-only tracking scenario

AUKF	IVKF	IPLF1(T)	IPLF5(T)	IPLF1	IPLF5
4	3	5	18	9	35

B. Range-bearings tracking scenario

We consider a 2-D scenario in which we measure range and bearings to show the benefits of the proposed filters when there are additional, non-directional measurements. In this case, we compare the proposed VMF-IPLF filters with the EKF, the UKF, the angular UKF [17], the truncated UKF (TUKF) [18] and the adaptive unscented Gaussian likelihood filter (AUGLAF) [33]. The last two filters can only be applied with measurement functions that have an inverse so they could not be applied in the previous scenario. The pseudolinear Kalman filter variants were developed for cases where there are only angular measurements, so they are not included here. Also, we have seen the superior performance of the proposed filters with respect to these filters in the previous example. Following [18], TUKF and UGLAF are implemented using circular sums and subtractions to deal with the angular measurement, as in [17]. The TUKF is implemented with parameter $\gamma = 0.1$, as in [18]. The unscented transform used in all filters that require sigma points is chosen as in Section VI-A.

As in Section VI-A, the state vector at time k is $x^k = \left[p_x^k, \dot{p}_x^k, p_y^k, \dot{p}_y^k\right]^T$. We consider the nearly-constant velocity model, which is given by (25), with the same parameters. The prior mean and covariance matrix are: $\bar{x}^0 = [5, 2, 5, 2]^T$ and $\Sigma^0 = \text{diag}\left(\left[20^2, 1, 20^2, 1\right]\right)$. The radar sensor is located at position $[0, 0]^T$. We analyse bearings measurements generated by a Gaussian distribution with variance $\sigma_\vartheta^2 = 1/\kappa$ and by a VMF distribution with concentration parameter κ. The range measurement z_r^k has additive Gaussian noise

$$z_r^k = \sqrt{\left(p_x^k\right)^2 + \left(p_y^k\right)^2} + \eta_r^k,$$

where η_r^k is a zero-mean Gaussian noise with variance σ_r^2.

We evaluate the algorithm with $N_{mc} = 10^3$ Monte Carlo runs with 50 different trajectories of length $K = 100$ drawn from the dynamic model, which has the same parameters of Section VI-A. We consider $\sigma_r^2 = 1$ and analyse the RMS error over all time steps, see (29) for different values of κ. As in the previous case, the EKF and UKF do not work well for Gaussian or VMF measurements, as sometimes the

Table II: RMS error over all time steps in 2-D bearings-only tracking scenario

	VMF measurement						Gaussian measurement					
κ_j	AUKF	IVKF	IPLF1(T)	IPLF5(T)	IPLF1	IPLF5	AUKF	IVKF	IPLF1(T)	IPLF5(T)	IPLF1	IPLF5
300	10.07	12.27	10.06	9.95	10.14	10.03	10.11	12.59	10.06	9.95	10.13	10.09
200	11.78	14.82	11.21	11.10	11.24	11.19	11.52	15.37	11.21	11.10	11.55	11.52
100	14.78	21.48	14.60	14.46	14.57	14.58	14.70	21.84	14.60	14.46	14.34	15.16

measurements are close to the $-\pi$ radians boundary, so they are not discussed further.

The resulting errors for the algorithms are shown in Table III. We can see that the best performing filter for all values of κ and VMF or Gaussian measurements is the proposed VMF-IPLF with 5 iterations. This filter is then followed by AUGLAF. TUKF and VMF-IPLF with Taylor series and 5 iterations. The overall ranking in performance ends with the VMF-IPLF1, VMF-IPLF1(T), and finally AUKF. In this scenario, in which there is a large uncertainty in the a prior position, it is highly useful to consider iterated filters, such as VMF-IPLF5 or VMF-IPLF5(T), or filters that invert the measurement function, such as AUGLAF or TUKF. Non-iterated Kalman-filter type algorithms, such as VMF-IPLF1, VMF-IPLF1(T) and AUKF perform worse. The execution times of the algorithms are shown in Table IV. The best performing filters have a higher computational burden.

C. Azimuth and elevation tracking scenario

In this section, we consider a 3-D target tracking scenario where the state vector at time k is $x^k = \left[p_x^k, \dot{p}_x^k, p_y^k, \dot{p}_y^k, p_z^k, \dot{p}_z^k\right]^T$, which includes position and velocity. The dynamic model of the target is the nearly-constant velocity model:

$$x^{k+1} = Fx^k + v^k, \tag{30}$$

$$F = I_3 \otimes \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix}, \tag{31}$$

where v^k is the process noise at time k. The process noise is zero-mean Gaussian distributed with covariance matrix

$$Q = qI_3 \otimes \begin{pmatrix} \tau^3/3 & \tau^2/2 \\ \tau^2/2 & \tau \end{pmatrix}, \tag{32}$$

where q is a parameter of the model. The prior at time 0 is

$$p(x^0) = \mathcal{N}\left(x^0; \bar{x}^0, \Sigma^0\right), \tag{33}$$

where

$$\bar{x}^0 = [-100, 7, 0, 5, 0, 0]^T,$$
$$\Sigma^0 = \text{diag}\left(\left[4^2, 1, 100^2, 1, 2^2, 0.01\right]\right).$$

In the simulations, we consider $q = 0.25$ and $\tau = 0.5$.

The target is observed by two sensors that measure bearing and elevation of the target at all time steps. The sensors are located at $[s_{x,1}, s_{y,1}, s_{z,1}]^T = [100, 0, 0]^T$ and $[s_{x,2}, s_{y,2}, s_{z,2}]^T = [-200, 0, 0]^T$. The measurements are distributed according to (5) with measurement function (6) extended to the 3-D case, and concentration parameter $\kappa_j = 600$. Gaussian distributed measurements will be analysed at the end of this subsection.

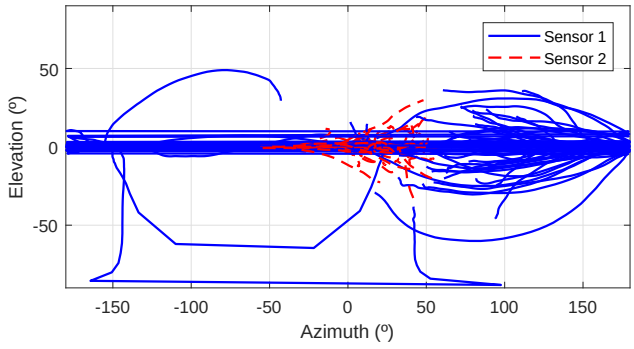

Figure 6: True azimuth and elevation of the considered trajectories with respect to the two sensors.

We have implemented the proposed VMF-IPLF with 1 and 5 iterations, using sigma-points and Taylor series, the conventional EKF and UKF, the AUKF for azimuth/elevation measurements [17], and the PLKF for azimuth/elevation measurements described in [21]. The filters that require an additive Gaussian measurement model use (23) and (24) with a zero-mean Gaussian noise with covariance matrix I_2/κ_j. We recall from Section V that this is an accurate approximation of the VMF measurement for low elevation.

We evaluate the algorithm with $N_{mc} = 10^3$ Monte Carlo runs with 50 different trajectories of length $K = 100$ drawn from the dynamic model. The true bearings and elevations of the 50 trajectories with respect to the two sensors are shown in Figure 6. It can be seen that the true azimuth for the first sensor crosses the π radians boundary, which makes the conventional EKF and UKF perform badly, so they are not analysed further. The RMS error for the position elements provided by the rest of the filters is shown in Figure 7. The best performing filters are VMF-IPLF5 and VMF-IPLF5(T) and they provide the lowest errors at practically all time steps. The second best performing filters are VMF-IPLF1 and VMF-IPLF1(T), which behave worse than the best filters at the beginning of the simulation. AUKF is the next filter in terms of performance. It roughly behaves as VMF-IPLF1 and VMF-IPLF1(T) at the beginning of the simulation but it performs worse than all the previous filters from around time steps 50 to 85. The worst performing filter is the PLKF. This is due to the fact that it makes the approximation that posterior distribution on the xy plane is independent of the distribution on the z axis, which is not accurate. As in the first scenario, the RMS error increases

Table III: RMS error over all time steps in range-bearings scenario

	VMF measurement							Gaussian measurement						
κ	AUKF	AUGLAF	TUKF	IPLF 1(T)	IPLF 5(T)	IPLF 1	IPLF 5	AUKF	AUGLAF	TUKF	IPLF 1(T)	IPLF 5(T)	IPLF 1	IPLF 5
300	4.79	3.22	3.34	6.51	3.43	4.27	3.12	4.78	3.20	3.35	6.51	3.41	4.27	3.11
200	5.08	3.70	3.88	6.96	3.89	4.55	3.57	5.09	3.76	3.93	6.98	3.88	4.58	3.60
100	5.77	4.85	5.11	7.92	4.89	5.26	4.59	5.83	4.93	5.18	7.98	4.89	5.34	4.63

Table IV: Execution times in milliseconds of the algorithms in range-bearings scenario

AUKF	AUGLAF	TUKF	IPLF1(T)	IPLF5(T)	IPLF	IPLF
7	23	23	6	23	11	47

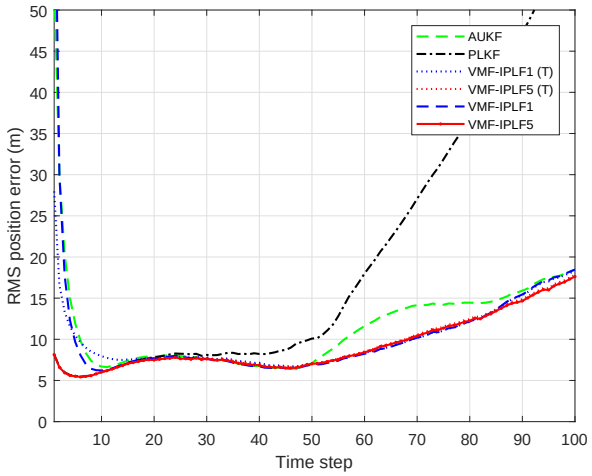

Figure 7: RMS position error against time for the 3-D scenario.

towards the end of the simulation as the target is moving away from the sensors, and the measurements are less informative in this case.

We also show the NEES for the position in Figure 8. VMF-IPLF5 and VMF-IPLF5(T) provide consistent estimates at all time steps. VMF-IPLF1 and VMF-IPLF1(T) are inconsistent at the beginning but they behave consistently afterwards. AUKF provides inconsistent estimates at many time steps. PLKF performs worse. The execution times of the algorithms are shown in Table V. Iterations increase the computational time but also improve performance.

Finally, we analyse the effect of parameter κ_j and the noise distribution (either VMF or Gaussian) on estimation performance. The Gaussian measurements according to (23) and (24) with a zero-mean Gaussian noise with covariance matrix I_2/κ_j. We show the RMS position error over all time steps, as defined in (29), in Table VI for $\kappa_j \in \{100, 300, 600\}$. For $\kappa_j \in \{300, 600\}$, the best performing filter is IPLF5 for VMF and Gaussian generated measurements, followed by IPLF5(T). Among the non-iterated filters IPLF1(T) is the one that performs best in general in this example and is the best

Table V: Execution times in milliseconds of the algorithms in 3-D scenario

AUKF	PLKF	IPLF1(T)	IPLF5(T)	IPLF1	IPLF5
11	12	14	61	18	83

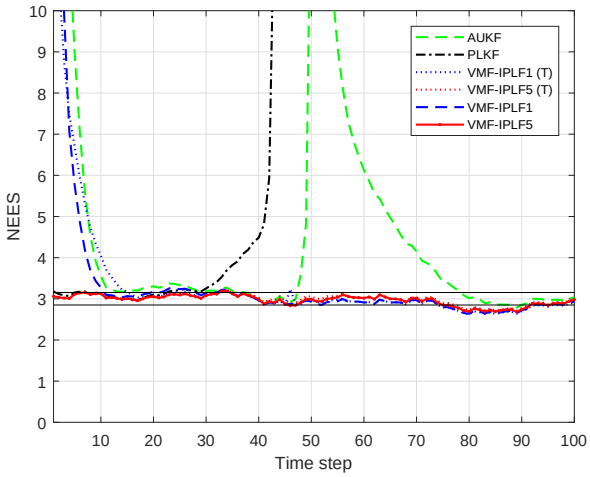

Figure 8: NEES of the position against time for the 3-D scenario. Two horizontal black lines indicate the 95% probability region.

filter for $\kappa_j = 100$. PLKF performs worse than the rest of the filters for the considered concentration parameters and distribution of the measurements.

VII. Conclusions

In this paper we have proposed a new Gaussian filter for target tracking with direction-of-arrival measurements. The proposed filter is derived from first principles using a von Mises-Fisher distribution for the angular measurements and deals in a unified way with 2-D and 3-D scenarios, which include bearings and azimuth/elevation measurements, respectively. The filter is based on performing statistical linear regression of the directional measurements and can be implemented using sigma-points or analytical linearisation. The linearisations include some correction factors due to the directional/von Mises-Fisher property of the measurements. In addition, iterations can be performed in each update with the aim to improve performance, especially, for informative measurements. Similarities and differences with nonlinear Kalman filters with additive Gaussian noise and angular measurements have been highlighted.

The proposed filters have been tested in three different scenarios involving direction-of-arrival measurements: bearings-only measurements, azimuth-elevation measurements, and range-bearings measurements. The proposed filters outperform Gaussian filters in the literature that deal with direction-of-arrival measurements.

Appendix A

In this appendix, we calculate the covariance matrix of the VMF distribution, see (1). We note that $\vartheta = \mu\kappa$ is the natural

Table VI: RMS error over all time steps in 3-D scenario

	VMF measurement						Gaussian measurement					
κ_j	AUKF	PLKF	IPLF1(T)	IPLF5(T)	IPLF1	IPLF5	AUKF	PLKF	IPLF1(T)	IPLF5(T)	IPLF1	IPLF5
600	13.06	26.29	10.80	10.00	12.96	9.87	11.89	19.97	10.83	10.19	13.02	10.05
300	14.52	35.94	13.05	12.89	14.87	12.69	13.83	28.95	13.11	12.99	14.99	12.77
100	18.88	55.04	17.84	19.18	19.20	17.98	18.24	49.59	18.06	19.52	19.39	18.71

parameter of the VMF distribution so, in this appendix, we write the VMF density as

$$\mathcal{V}(z;\vartheta) = \exp\left(\vartheta^T z - \log\rho(\|\vartheta\|)\right)\chi_{\|z\|=1}(z),$$

where the normalising constant is

$$\rho(\|\vartheta\|) = \frac{\Gamma(n/2)\,\mathrm{I}_{n/2-1}(\|\vartheta\|)}{\left(\frac{\|\vartheta\|}{2}\right)^{n/2-1}},$$

and we recall that $\mathrm{I}_a(\cdot)$ was defined in (1). By exponential family theory, we have that [34]

$$\mathrm{E}[z] = \nabla_\vartheta \log\rho(\|\vartheta\|),$$
$$\mathrm{C}[z] = \nabla^2_\vartheta \log\rho(\|\vartheta\|).$$

Using (2), we have

$$\nabla_\vartheta \log\rho(\|\vartheta\|) = A_n(\|\vartheta\|)\frac{\vartheta}{\|\vartheta\|}.$$

Computing the Hessian, we obtain

$$\nabla^2_\vartheta \log\rho(\|\vartheta\|) = \frac{A_n(\|\vartheta\|)}{\|\vartheta\|}I_n + \left[A_n'(\|\vartheta\|) - \frac{A_n(\|\vartheta\|)}{\|\vartheta\|}\right]\frac{\vartheta}{\|\vartheta\|}\frac{\vartheta^T}{\|\vartheta\|}.$$

We now use the equivalence [15, Eq. (10.3.17)]

$$A_n'(\|\vartheta\|) = 1 - A_n^2(\|\vartheta\|) - \frac{n-1}{\|\vartheta\|}A_n(\|\vartheta\|),$$

and $\vartheta = \mu\kappa$ to yield (3).

APPENDIX B

In this appendix, we show that for $\kappa_j \to 0$ for $j \in \{1,...,m\}$, the posterior and the approximated posterior tend to the prior.

A. Posterior

For $\kappa_j \to 0$, the VMF distribution tends to the uniform distribution on the sphere [15], which yields

$$\lim_{\kappa_j\to 0}\mathcal{V}(z_j; h_j(x), \kappa_j) = \chi_{\|z\|=1}(z),$$

where we recall that the VMF densities in this paper are given with respect to the uniform distribution.

This result implies that the measurement does not depend on the state for $\kappa_j \to 0$, $j \in \{1,...,m\}$, so the posterior is equal to the prior.

B. Approximated posterior

We compute the density of the approximated posterior for $\kappa_j \to 0$ for $j \in \{1,...,m\}$. For small κ_j and using [15, Eq. (10.3.9)], we obtain

$$\frac{A_n(\kappa_j)}{\kappa_j} = \frac{1}{n}\left[1 - \frac{\kappa_j^2}{n(n+2)} + \frac{\kappa_j^4}{n^2(n+2)(n+4)} + O\left(\kappa_j^6\right)\right].$$

Using this equation, we obtain the following limits

$$\lim_{\kappa_j\to 0} A_n(\kappa_j) = 0,$$
$$\lim_{\kappa_j\to 0}\frac{A_n(\kappa_j)}{\kappa_j} = 1/n.$$

Substituting these results into (18)-(21), we obtain

$$\mathrm{E}[g_j(x)] = 0, \qquad \mathrm{C}[g_j(x)] = 0,$$
$$\mathrm{C}[x, g_j(x)] = 0, \qquad \mathrm{E}[R_j(x)] = \frac{1}{n}I_n,$$

which implies that $A_j^+ = 0$, $b_j^+ = 0$ and $\Omega_j^+ = \frac{1}{n}I_n$. Then, the updated mean and covariance, see (11) and (12), are $\overline{u} = \overline{x}$ and $W = P$, which implies that the filter output coincides with the prior, which coincides with the posterior.

APPENDIX C

In this appendix, we prove that the update based on a (nonlinear) Kalman filter with angular measurement and additive Gaussian noise, see (23), is approximately equal to the update of one iteration of the proposed filter with a 2-D VMF measurement model under the following conditions

1) The density of the predicted angular measurement is narrowly concentrated around a point and is negligible around the π radians boundary.
2) The variance of the angular measurement noise used in the Kalman filter is $R_\varphi = \frac{1}{\kappa}$.

Note that condition 1) implicitly implies that the concentration parameter is κ is high. In this appendix, we denote the angle of measurement z as y, which implies that $z = [\cos y, \sin y]^T$. It should be noted that if z has a von Mises distribution on the circle, which corresponds to the 2-D VMF distribution, with mode $\mu = [\cos m, \sin m]$ and concentration parameter κ, then y has a von Mises distribution with mode m and concentration parameter κ, which has a density on the angle given by [15]

$$\mathcal{V}_a(y; m, \kappa) = \frac{\exp(\kappa\cos(y-m))}{2\pi\mathrm{I}_0(\kappa)}. \tag{34}$$

For high κ, one has [15]

$$\mathcal{V}_a(y; m, \kappa) \approx \mathcal{N}\left(y; m, \frac{1}{\kappa}\right).$$

This implies that if the conditional density of the measurement z is

$$p\left(z\left|x\right.\right)=\mathcal{V}\left(z;h\left(x\right),\kappa\right),$$

with $h\left(x\right)$ given by (6), the corresponding distribution on y is

$$p\left(y\left|x\right.\right)=\mathcal{V}_{a}\left(z;h_{a}\left(x\right),\kappa\right),$$

where $h_{a}\left(x\right)=\operatorname{atan2}\left(p_{y}-s_{y,j},p_{x}-s_{x,j}\right)$. For high κ (condition 1), we have

$$p\left(y\left|x\right.\right)\approx\mathcal{N}\left(y;h_{a}\left(x\right),R_{\varphi}\right),\tag{35}$$

where the variance of the noise is $R_{\varphi}=\frac{1}{\kappa}$ (condition 2). This proves the equivalence between von Mises measurement models and Gaussian additive models for high κ and one angle. Nevertheless, the conventional Kalman filter update operates on the angle y, while the proposed filter operates on the directional vector z. In the following, we prove that the two approaches, considering one iteration of the proposed filter, are approximately equal under the conditions stated above.

If we only consider one iteration, the updated mean and covariance of the VMF filter are given by

$$\overline{u}=\overline{x}+\mathrm{C}\left[x,z\right]\left(\mathrm{C}\left[z\right]\right)^{-1}\left(z-\mathrm{E}\left[z\right]\right),\tag{36}$$

$$W=P-\mathrm{C}\left[x,z\right]\left(\mathrm{C}\left[z\right]\right)^{-1}\left(\mathrm{C}\left[x,z\right]\right)^{T}.\tag{37}$$

We proceed to compute the required moments of z in terms of the moments of y to show the equivalence.

As z is a function of variable y, the moments of z in relation to the moments of y are

$$\mathrm{E}\left[z\right]=\left[\begin{array}{cc}\mathrm{E}\left[\cos y\right] & \mathrm{E}\left[\sin y\right]\end{array}\right]^{T},\tag{38}$$

$$\mathrm{C}\left[x,z\right]=\left[\begin{array}{cc}\mathrm{C}\left[x,\cos y\right] & \mathrm{C}\left[x,\sin y\right]\end{array}\right],$$

$$\mathrm{C}\left[z\right]=\left[\begin{array}{cc}C_{11} & C_{12}\\ C_{12} & C_{22}\end{array}\right],\tag{39}$$

where

$$C_{11}=\mathrm{E}\left[\cos^{2}y\right]-\mathrm{E}\left[\cos y\right]^{2},$$

$$C_{12}=\mathrm{E}\left[\cos y\sin y\right]-\mathrm{E}\left[\cos y\right]\mathrm{E}\left[\sin y\right],$$

$$C_{22}=\mathrm{E}\left[\sin^{2}y\right]-\mathrm{E}\left[\sin y\right]^{2}.$$

We now note that the origin of the angular measurements, that is, the direction the 0 radians is pointing at can be changed, due to the symmetry of the measurements. In other words, adding any offset angle to the measurements (and measurement model) does not affect the Bayesian update or Kalman filter updates, in the latter case, under condition 1. Then, without loss of generality, we consider that the origin of the angular measurements is chosen such that $\mathrm{E}\left[y\right]=0$.

Under condition 1, we can then make the small angle approximations

$$\cos y\approx1,\,\sin y\approx y,\tag{40}$$

based on a a first-order Taylor series linearisation around $y=0$, to approximate $\mathrm{E}\left[z\right]$ and $\mathrm{C}\left[x,z\right]$. In this case, we obtain

$$\mathrm{E}\left[z\right]\approx\left[\begin{array}{cc}1 & \mathrm{E}\left[y\right]\end{array}\right]^{T},\tag{41}$$

$$\mathrm{C}\left[x,z\right]\approx\left[\begin{array}{cc}0 & \mathrm{C}\left[x,y\right]\end{array}\right].\tag{42}$$

If we use (40) to approximate $\mathrm{C}\left[z\right]$, the resulting matrix is not invertible. Therefore, we require a higher accuracy to approximate this term and we consider the Taylor series up to the second order terms of the trigonometric functions in $\mathrm{C}\left[z\right]$. That is,

$$\cos y\approx1-y^{2}/2,\quad\sin y\approx y,\tag{43}$$

and

$$\cos^{2}y\approx1-y^{2},\quad\sin^{2}y\approx y^{2},\quad\cos y\sin y\approx y.\tag{44}$$

We substitute (43) and (44) into (39) to obtain

$$\mathrm{C}\left[z\right]\approx\left[\begin{array}{cc}-\frac{\mathrm{E}\left[y^{2}\right]^{2}}{4} & 0\\ 0 & \mathrm{E}\left[y^{2}\right]\end{array}\right].\tag{45}$$

Then, substituting (41), (42), (45), and $z=\left[\cos y,\sin y\right]^{T}\approx\left[1,y\right]^{T}$ into (36), we have

$$\begin{aligned}\overline{u}&\approx\overline{x}+\left[\begin{array}{cc}0 & \mathrm{C}\left[x,y\right]\end{array}\right]\left[\begin{array}{cc}-\frac{4}{\mathrm{E}\left[y^{2}\right]^{2}} & 0\\ 0 & \frac{1}{\mathrm{E}\left[y^{2}\right]}\end{array}\right]\left[\begin{array}{c}1-1\\ y-0\end{array}\right]\\&=\overline{x}+\frac{\mathrm{C}\left[x,y\right]}{\mathrm{C}\left[y\right]}y.,\\&=\overline{x}+\frac{\mathrm{C}\left[x,h_{a}\left(x\right)\right]}{\mathrm{C}\left[h_{a}\left(x\right)\right]+R_{\varphi}}y.\end{aligned}\tag{46}$$

where we have used (35) and $\mathrm{E}\left[y\right]=0$. Equivalently, for the covariance matrix (37), we obtain

$$W\approx P-\frac{\mathrm{C}\left[x,h_{a}\left(x\right)\right]}{\mathrm{C}\left[h_{a}\left(x\right)\right]+R_{\varphi}}\left(\mathrm{C}\left[x,h_{a}\left(x\right)\right]\right)^{T}.\tag{47}$$

Equations (46) and (47) correspond to the KF update equations when we use the angle as the measurement and the average angle $\mathrm{E}\left[y\right]=0$, as considered. Therefore, this proves that, under conditions 1)-2), the update based on one iteration of the proposed VMF filter is equivalent to the conventional Kalman filter update with angular measurements and additive noise.

REFERENCES

[1] Y. Bar-Shalom, T. Kirubarajan, and X. R. Li, *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., 2001.

[2] D. Dardari, P. Closas, and P. M. Djuric, "Indoor tracking: Theory, methods, and technologies," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 4, pp. 1263–1278, April 2015.

[3] K. Radnosrati, F. Gunnarsson, and F. Gustafsson, "New trends in radio network positioning," in *18th International Conference on Information Fusion*, July 2015, pp. 492–498.

[4] Y. He, A. Behnad, and X. Wang, "Accuracy analysis of the two-reference-node angle-of-arrival localization system," *IEEE Wireless Communications Letters*, vol. 4, no. 3, pp. 329–332, June 2015.

[5] A. Tahat, G. Kaddoum, S. Yousefi, S. Valaee, and F. Gagnon, "A look at the recent wireless positioning techniques with a focus on algorithms for moving receivers," *IEEE Access*, vol. 4, pp. 6652–6680, 2016.

[6] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

[7] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[8] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[9] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, June 2009.

[10] A. Farina, "Target tracking with bearings-only measurements," *Signal Processing*, vol. 78, pp. 61–78, Oct. 1999.

[11] B. Ristic and M. S. Arulampalam, "Tracking a manoeuvring target using angle-only measurements: algorithms and performance," *Signal Processing*, vol. 83, no. 6, pp. 1223–1238, June 2003.

[12] D. C. Chang and M. W. Fang, "Bearing-only maneuvering mobile tracking with nonlinear filtering algorithms in wireless sensor networks," *IEEE Systems Journal*, vol. 8, no. 1, pp. 160–170, March 2014.

[13] A. Giannitrapani, N. Ceccarelli, F. Scortecci, and A. Garulli, "Comparison of EKF and UKF for spacecraft localization via angle measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 1, pp. 75–84, Jan. 2011.

[14] S. Challa, M. R. Morelande, D. Musicki, and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge University Press, 2011.

[15] K. V. Mardia and P. E. Jupp, *Directional Statistics*. John Wiley & Sons, 2000.

[16] S. R. Jammalamadaka and A. SenGupta, *Topics in Circular Statistics*. World Scientific, 2001.

[17] D. F. Crouse, "Cubature/unscented/sigma point Kalman filtering with angular measurement models," in *18th International Conference on Information Fusion*, July 2015, pp. 1550–1557.

[18] A. F. García-Fernández, M. R. Morelande, and J. Grajal, "Truncated unscented Kalman filtering," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3372–3386, July 2012.

[19] V. J. Aidala, "Kalman filter behavior in bearings-only tracking applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 15, no. 1, pp. 29–39, Jan. 1979.

[20] N. H. Nguyen and K. Doğançay, "Improved pseudolinear Kalman filter algorithms for bearings-only target tracking," *IEEE Transactions on Signal Processing*, vol. 65, no. 23, pp. 6119–6134, Dec. 2017.

[21] S. Xu, K. Doğançay, and H. Hmam, "3D pseudolinear Kalman filter with own-ship path optimization for AOA target tracking," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2016, pp. 3136–3140.

[22] N. H. Nguyen and K. Doğançay, "Instrumental variable based Kalman filter algorithm for three-dimensional AOA target tracking," *IEEE Signal Processing Letters*, vol. 25, no. 10, pp. 1605–1609, Oct. 2018.

[23] I. Arasaratnam, S. Haykin, and R. Elliott, "Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 953–977, May 2007.

[24] F. Tronarp, A. F. García-Fernández, and S. Särkkä, "Iterative filtering and smoothing in non-linear and non-Gaussian systems using conditional moments," *IEEE Signal Processing Letters*, vol. 25, no. 3, pp. 408–412, March 2018.

[25] K. V. Mardia and S. A. M. El-Atoum, "Bayesian inference for the von Mises-Fisher distribution," vol. 63, no. 1, pp. 203–206, Apr. 1976.

[26] G. S. Watson, *Statistics on Spheres. University of Arkansas Lecture Notes in the Mathematical Sciences. Vol. 6.* John Wiley & Sons, 1983.

[27] B. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice-Hall, 1979.

[28] A. F. García-Fernández, L. Svensson, M. R. Morelande, and S. Särkkä, "Posterior linearization filter: principles and implementation using sigma points," *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5561–5573, Oct. 2015.

[29] M. R. Morelande and A. F. García-Fernández, "Analysis of Kalman filter approximations for nonlinear measurements," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5477–5484, Nov. 2013.

[30] M. Skolnik, *Introduction to Radar Systems*. McGraw-Hill, 2001.

[31] P. Berens, "CircStat: A MATLAB toolbox for circular statistics," *Journal of Statistical Software*, vol. 31, pp. 1–21, Sep. 2009.

[32] Y. H. Chen, D. Wei, G. Newstadt, M. DeGraef, J. Simmons, and A. Hero, "Parameter estimation in spherical symmetry groups," *IEEE Signal Processing Letters*, vol. 22, no. 8, pp. 1152–1155, Aug. 2015.

[33] A. F. García-Fernández, M. R. Morelande, J. Grajal, and L. Svensson, "Adaptive unscented Gaussian likelihood approximation filter," *Automatica*, vol. 54, pp. 166–175, April 2015.

[34] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.